A fast, easy, and efficient estimator for the trade flux between heterogeneous economies

Robert Sova and Anamaria Sova

Academia de Studii Economice (Academy of Economic Studies)

5. March 2007

A FAST, EASY, AND EFFICIENT ESTIMATOR FOR THE TRADE FLUX BETWEEN HETEROGENEOUS ECONOMIES

Robert ŞOVA,
Academy of Economic Studies of Bucharest[1]
Anamaria ŞOVA,
Academy of Economic Studies of Bucharest[2]

Abstract

Compared to time-series or cross-section analyses, panel data allow us to control for individual specific characteristics - possibly unobservable - which may be correlated with certain explanatory variables in the specification of an economic relationship. Not controlling for unobservables leads to obtaining biased results. After controlling for such unobservable characteristics, we calculate efficient estimates of a trade flux equation between heterogeneous economies.

Key words: international trade, POLS estimators, individual heterogeneity, fixed effects, random effects

1. INTRODUCTION

The fundamental advantage of a panel data set over a cross section is that it will allow the researcher great flexibility in modeling differences in behavior across individuals [Green, 2002]. In empirical analysis of data consisting of repeated observations on economic units it is often assumed that we have observable random variables, and unobserved effect. The unobserved effect is often interpreted as capturing features of an individual. In fact, a primary motivation for using panel data is to solve the omitted variables problem [Wooldridge, 2002].

One seeks has to test the null hypothesis of the unobserved characteristics presence. If the null assumption is rejected, then one must include individual effects in the model. In the traditional approach to panel data models, unobserved effect, is called a „random effect" when it is treated as a random variable and a „fixed effect" when it is treated as a parameter to be estimated for each cross section observation. We follow Wooldridge (2002) in the sense that unobserved effects are treated as random variables, and the key issue is whether the unobserved effect is uncorrelated with the explanatory variables. A minimum condition, in the linear panel data models is that the error in each time period was assumed to be uncorrelated with the explanatory variables in the same time period but for certain panel data applications this assumption is too strong. Basically we has two alternative different estimation methods "random effects" estimation (RE) which is associated with GLS estimator and "fixed effects" estimation (FE) with „within" estimator. To identify whether

1 Piata Romana nr. 6, tel. 00 40 21 319 19 89, e-mail robertsova@yahoo.com
2 Piata Romana nr. 6, tel. 00 40 21 319 19 89, e-mail anamariasova@yahoo.fr

the unobservables are correlated with explanatory factors, we perform a Hausman test comparing the fixed effects and random effects estimators. The test is based on the fact that the random effects estimator is biased if unobservables are correlated with the explanatory variables, while the fixed effects estimator is always unbiased but is less efficient if there is no correlation. The gain in efficiency results from the utilization of the „between" estimator in addition to the „within" estimator. Furthermore, when the effects are not correlated with the explanatory variables, the within and between estimators are the same and therefore any weighted matrix combination thereof will be the same [Mundlak 1978].

Actually, bilateral trade can be influenced by specific characteristics. For instance, the impacts of historical, cultural and political links on trade flows are difficult to observe and quantify.

The choice of the method (FE or RE) depends on two important things, its economic and econometric relevance. From an economic point of view there are unobservable time invariant random variables, difficult to be quantified, which may simultaneously influence some explanatory variables and the trade volume. From an econometric point of view the inclusion of fixed effects is preferable to random effects because the rejection of the null hypothesis of correlation of the unobservable characteristics with some explanatory variables is less plausible (Baier and Bergstrand 2005).

We now briefly present the panel data econometric methods used in our paper to estimate the possible various specifications of our models: pooled OLS estimator (POLS), random effect estimator (RE), within estimator (FE).

Estimating the trade flux between heterogeneous economies

The bilateral trade can be influenced by specific characteristics. For instance, the impacts of historical, cultural and political links on the trade flows between heterogeneous economies are difficult to observe and quantify.

Our model proposed to estimation is a two dimension gravity model as the following:

$$\mathbf{Log(X_{ijt})=a_0 + a_1 log(GDP_{it}) + a_2 log(GDP_{jt}) + a_3 log(DGDPC_{ijt}) + a_4 log(Dist_{ij}) + a_5 log(Tchr_{ijt}) + a_6 Acc_{ijt} + u_{ij} + d_t + \varepsilon_{ijt}} \quad (1)$$

for (i = 1,……N; t = 1,…..T)
where X_{ijt} denotes the bilateral trade between countries i and j at time t with $i \neq j$ (source : CHELEM – CEPII data base), ao is the intercept, GDP_{it}, GDP_{jt} represents the Gross Domestic Product of country i and country j at time t (source : CHELEM CEPII – data base), $DGDPC_{ijt}$ is the difference of GDP per capita between partners and is a proxy of economic distance or of comparative advantage intensity, $Dist_{ij}$ represents the distance between two countries (CEPII data base), proxy variable for costs of transport, $Tcrh_{ijt}$ is the real exchange rate which indicates the competitiveness of price, Acc_{ijt} is a dummy variable

that equals 1 if country i and country j have signed a regional agreement, and zero otherwise, u_{ij} is the unobserved bilateral effects (i = 1,2,…,N, j=1,2,…,M), d_t time specific effect, The time effects account for the business cycle and changes in openness across all countries t = 1,…..T and εijt idiosyncratic errors assumed to have a normally distribution with zero mean and constant variance for all observations and pair wise uncorrelated

The estimation period goes from 1990 to 2004, i.e. 15 years for a sample of 19 developed countries (OECD) and 2 developing countries[3] (CEEC - Central and Eastern European countries). Data are organized in panel with two dimensions: country pairs, and years.

We compared various different method of estimation as POLS, FE, RE to study the effect of integration on trade flux performances. Random effects and fixed effects are still the most popular approaches to estimating unobserved effects panel data models under strict exogeneity of the explanatory variables. If the coefficients of the fixed model different systematic from those are estimated by the random model can be interpreted, following Mudlak (1978), that there is a skew of correlation between some of the explanatory variables and the bilateral specific effect.

The results of POLS, FE, RE estimations are reported in the table 1 in appendix.

The null hypothesis of no existence of individual effects is rejected. The F statistic test is 40.7. Thus, POLS not controlling for heterogeneity run the risk of obtaining biased results. The comparison between columns (2) and (3) shows several differences concerning the estimated coefficients of time-varying explanatory variables. The HAUSMAN test of the difference between random effects and within estimates confirms this result, but the difference in coefficients was not systematic. The null hypothesis of no correlation is accepted. The test statistic is 18.63, which is distributed as $Chi^2 1$ with $Prob>Chi^2$ 0.48.

A comparison between the three estimation leads to conclusion that using POLS, we cannot expect to consistently estimate any parameters of the model when the omitted variable is ignore. With FE estimator we are obtained unbiased results but less efficient, and with the cost of losing the possibility to determine the influence of the time-invariant variable. The results indicate no correlation existence between the unobservable effect and the explanatory variables, and so RE estimator are unbiased and efficient.

The difference of income per capita has the expected positive sign which is in accordance with the Heckscher-Ohlin theory, i.e. the trade between two zones is based on comparative advantage. It's a complementary inter-industry trade where less developed countries are specialized in labor intensive industries and where the wage costs are less expensive. The variables like country size which have the most important coefficients explain better the level of bilateral exchanges. The international organization membership has a low positive

3 Romania, Bulgaria

influence (0.297) on trade flows. On the contrary, the distance variable represents an obstacle for trade. It should be noted that the distance between countries have an important elasticity (-1.283) and hence have an important explanatory capacity.

4. CONCLUSION

The use of panel econometric method in empirical analysis of trade flows is convenient because it permits for controlling the individual heterogeneity to avoid biased results. As it is known, the time-series and cross-section not controlling for heterogeneity run the risk of obtaining biased results[4].

From the analysis of trade flows estimates between heterogeneous economies, more particularly from the trade between two CEEC – the last integrating countries in UE - and OECD countries with different levels of economic development one can draw several conclusions:

From an econometric point of view the use of an RE estimator for the gravity model appears convenient for our data sample.

In our analysis the time invariant variable "geographical distance" (-1.283) and the country sizes variables ($GDP_j = 0.973$; $GDP_i = 0.940$) have a major importance in international trade flows explanation.

From an economic point of view, trade flows developed between CEEC-2 and OECD countries with heterogeneous economies and different levels of economic development are inter-industry trade and vertical intra-industry trade.

References

[1] Baier, L.S., Bergstrand, J.H, "Do Free Trade Agreements Actually increase Members' International Trade?", FRB of Atlanta Working Paper No. 2005-3., 2005

[2] Baltagi B.H., "Econometric Analysis of Panel Data", John Wiley & Sons Ltd, New York, 2nd edition.,2001

[3] Carrere C.,"Revisiting the Effects of Regional Trading Agreements on Trade Flows with Proper Specification of Gravity Model", European Economic Review vol. 50, 223-247, 2006

[4] Cheng I.-H.,Wall, H. Controlling for heterogeneity in gravity models of trade and integration, Federal Reserve Bank of Saint Louis Review 87, 49-63., 2005

[5] Egger P., "A note on the Proper Econometric Specification of the Gravity Equation", Economics Letters, Vol. 66, pp.25-31,2000

[6] Egger,P., "An Econometric View on the Estimation of Gravity Models and the Calculation of Trade Potentials", The World Economy 25 (2), 297 – 312,2002

[7] William H. Greene, "Econometric Analysis" Fifth Edition, Pearson Education, 2002

[4] Badi H. Baltagi (2001)

[8] Hausman J.A. and Taylor,W.E.,1981, “Panel Data and Unobservable Individual Effects”,Écometrica, 49, pp. 1377-1398,1981.
[9] Helpman, E. and Krugman, P., “Market Structure and Foreign Trade. Increasing Returns, Imperfect Competition, and the International Economy”, Cambridge MA/ London: MIT Pres. 1985.
[10] Linnemann, H., “An Econometric Study of International Trade Flows”, North Holland Publishing Company, Amsterdam, 1966.
[11] Mundlak Y., “On the Pooling of Time Series and Cross Section”, Econometrica, vol. 46, no. 1, pp 69-85, 1978
[12] Sapir A., Winter C., “Services Trade”, Surveys in International Trade, edited by David Greenaway and L. Alan Winters, Oxford, UK: Blackwell, 1994.
[13] Sova A., Sova, R., “Le commerce international dans les conditions de l’intégration européenne », Romanian Statistical Review, No 12. p.79-86, 2006
[14] Wooldrige, J.H., “Econometric Analysis of Cross Section and Panel Data 2rd Edition”: Books: The MIT Press Cambridge, Massachusetts London, England,, 2002.

Appendix

Table 1

VARIABLES	OLS	WITHIN	RANDOM
	(1)	(2)	(3)
	Xij	Xij	Xij
GDP_i	0.927	0.954	0.940
	(43.19)***	(4.47)***	(13.51)***
GDP_j	0.891	1.664	0.973
	(41.52)***	(7.81)***	(13.98)***
$Dist_{ij}$	-1.298	0.000	-1.283
	(-35.94)***	(.)	(-12.13)***
$DGDPC_{ij}$	0.219	0.130	0.157
	(2.65)***	(1.24)*	(2.10)**
$Tchr_{ij}$	-0.027	-0.028	-0.003
	(-1.95)*	(-1.91)*	(-1.98)*
Acc_{ij}	0.232	0.299	0.297
	(6.54)***	(12.47)***	(12.44)***
D_time	***	***	***
Constant	-4.114	-12.303	-4.455
	(10.97)***	(6.54)***	(7.03)***
Observations	1140	1140	1140
R-squared	0.78	0.69	0.78
Number of cod_rel		76	76
VIF	1.97		
F test that all α_i=0:	40.70 (0.00)	-	-
Ramsey-Reset Prob>Chi2	24.31 (0.00)	-	-
Breusch - Pagan / Cook – Weisberg (before correction)	117.76 (0.00)	-	-
Hausman test Prob>Chi^2_1	-	-	18.63 (0.48)
Absolute value of t statistics in parentheses			
* significant at 10%; ** significant at 5%; *** significant at 1%			